CHALMERS

Valanalys

Ett Twitter-baserat analysverktyg

Valanlys.se: A Graphical Trend Analysis Tool

Visualizing Political Trends From Semantically Analyzed Twitter Posts

Master of Science Thesis in the Programme Interaction Design

NIKLAS VON TELL
FREDRIK JÄDERSTRÖM SCHERSTÉN

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg, Sweden, June 2010

Valanlys.se: A Graphical Trend Analysis Tool
Visualizing Political Trends From Semantically Analyzed Twitter Posts

NIKLAS VON TELL
FREDRIK JÄDERSTRÖM SCHERSTÉN

Examiner: MORTEN FJELD

Department of Computer Science and Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Department of Computer Science and Engineering
Göteborg, Sweden June 2010

Abstract

The Internet has during a relatively short period of time changed people's perception on information and its availability. It has never before been so easy to get information about practically anything, just a few clicks away. The problem today is that we are living in a society with an information overflow, it can be really time consuming finding relevant facts in the ocean of data. A young entrepreneurial company, Saplo AB, thought of a solution: let the computers find whats relevant for us and save us both time and effort. They created a highly advanced algorithm that works in a similar fashion to the human brain. With the algorithm computers can now read, understand and sort text. Saplo has so far only focused their services toward other companies but now they want to show the rest of the world that their technique is capable of. This master thesis aims to provide Saplo with a platform for information visualization, a tool to visually present the result of their technology. To narrow down the work, the Swedish election was chosen as the subject for this master thesis. What it tries to answer is: "How to create a real-time feedback system for the political environment with automated trend-analysis?"

The master thesis goes through all necessary steps to complete this visualization tool from research to implementation and testing. The project results in a fully functional website that gathers tweets (messages) from Twitter that are related to the Swedish election. All the tweets are sent to Saplo to be analyzed and are there given a trend index. They are then forwarded to the website and presented in real-time. The trend index is used to draw a trend for the political party that the tweet is concerning. The master thesis also discusses possible improvements and interesting ideas for further development.

Sammanfattning

Internet har på relativt kort tid förändrat vår syn på information och dess tillgänglighet. Det har aldrig tidigare varit så enkelt att få tag på information om praktiskt taget vad som helst, bara några mus-klick bort. Problemet är snarare att vi i dagens samhälle drunknar i för mycket information, tiden läggs istället på att sortera ut relevant information. Ett ungt entreprenörsmässigt företag, Saplo AB, har kommit på en lösning: låta datorerna hitta vad som är relevant information och spara oss både tid och ansträngning. Saplo skapade en högst komplex algoritm som arbetar på ett liknande sätt som den mänskliga hjärnan. Med denna algoritm kan datorer nu läsa, förstå och sortera text för oss. Saplo har än så länge endast fokuserat sin service mot andra företag men nu ville de ta steget att visa världen vad deras teknik är kapabel till. Examensarbetet försöker tillhandahålla Saplo en plattform för informations visualisering, ett verktyg för att visuellt presentera resultatet av deras teknologi. För att begränsa arbetet till en rimlig nivå, valdes det svenska valet som ämne för detta examens arbetet. Examensarbetet vill besvara frågeställningen: "Hur skapar man ett real-tids feedback system för politiska klimatet med automatiska trendanalyser?".

Vi går igenom alla nödvändiga faser i utvecklingen, från undersökning till implementering och testning. Projektet resulterade i en fullt fungerande hemsida som samlar in tweets (meddelanden) ifrån Twitter som är relevanta till det Svenska valet. Alla tweets skickas till Saplo för att bli analyserade och ges där ett trendindex. Därefter skickas de vidare till hemsidan för publicering i realtid. Tweetets tillhörande trendindex används för att rita upp en trend för det parti som tweetet nämnde. Examensarbetet kommer även ta upp en diskussion angående förbättringar och förslag på framtida förändringar.

Preface

This Master Thesis is the final work at the M. Sc. Programme in Interaction Design at the Department of Computer Science and Engineering at Chalmers University of Technology. The thesis consists of 30 credits, which is equivalent to 20 weeks of work. The thesis and work was carried out for Saplo AB in Malmö, Sweden. This thesis is a part of a project for Saplo AB which will gather Twitter posts regarding the Swedish election and analyze the sentiment of the texts. The tweets will be presented together with the sentiment trends in a web application developed for this purpose and will serve as a way to gain publicity for the company and their services.

We want to give a special thanks to all the people who helped us at Saplo: Mattias Tyrberg, Fredrik Hörte, Anders Hall, Joakim Stenberg.

Contents

Chapter 1

Introduction

The amount of data on the Internet is constantly growing and thus making it more and more difficult to find, overview and handle what information is valuable. Saplo AB has developed an algorithm to automatically analyze large quantities of text without the traditional use of "keywords". Their algorithms are developed to imitate the human brain and should supposedly be more accurate and human-like in their analysis than their competitors. This technology has of course drawn a lot of attention to the company and Saplo was recently ranked as one of the most promising start-up companies in Europe.

A trend that is becoming clearer is that companies aim to provide easier and more user-friendly interfaces instead of a lot functionality. Saplo AB is very aware of this trend and therefore aims their products in this manner. The master thesis aims to provide Saplo with a tool for information visualization in an attractive yet informative way.

1.1 Background

Saplo AB is a young innovative company that has specialized in semantic text analysis. They provide three main services: Sentiment Analysis, Semantic Matching and Semantic Tags. This type of technology has great potential as it could save companies both money and labor. This might be the reason why Saplo has been given so much attention for their award winning ideas. Saplo's algorithms don't use specific keywords to analyze text in contrast to many of its competitors. Saplo's technique is therefore very unique and supposedly more accurate than its rivals.

Saplo has so far focused on marketing themselves to companies, while developing their product, and not the public domain. Up until now they have had neither the incentive nor the resources to target the public. They now however feel that they have the opportunity to develop a product with the sole purpose of promoting their technology and brand name.

After dialogs with Saplo AB we came to the conclusion to create a component based trend-visualizing tool concerning the Swedish election. This would be a great subject due to its promotional value and its relevance for the Swedish population. However, Saplo requested that all the components should be dynamic and applicable in other situations as well. Saplo wanted to aim the project to Twitter due to its increasing popularity in Sweden, although there are limitations in Twitter that Saplo needs to consider. One of these are that Twitter limits all their posts to 140 characters which can make it harder for Saplo's algorithms to analyze and present an accurate result.

The idea is that twitter posts, which are known as tweets, will be presented in real-time in a manner showing how positive or negative they are towards certain elements of the election. Saplo also put emphasis on using a line chart to plot the trends. These are requests that we need to take into consideration when designing the interface. They also requested the application to be very flexible in the sense that it should be possible to use on a wide variety of platforms. The application should also be able to be viewed on a TV monitor.

There are already some applications on the market, which perform subsets of similar services. They however make use of keywords giving less reliable results than Saplo's algorithms. An important part of this task is information visualization, presenting the analysis in an interesting and appealing way.

Social media can play a very important part in politics and by understanding this is a great advantage for the politicians, take for example the case of the US presidential election in 2008. Here presidential nominee Barack Obama used new media to reach out to new groups of people, especially younger first time voters [8].

So to summarize, the objective of the master thesis is to create an analysis tool, to follow political-trends during the Swedish election using information from Twitter. The product should however be developed in a much more dynamical matter as Saplo wants this product to be so flexible that it could be applied on almost any topic. The master thesis will lay the foundation to potential further developing the product so that the user can enter a keyword in the application and Twitter posts (tweets) containing the specified keyword will be analysed by Saplo's software.

1.2 Goals

Politics is perceived by most people to be rather complicated and unexciting. There could be many reasons why some people feel this way. We believe that one of the reasons is due to the lack of influence and tools to help them see how media and other people actually can affect the politicians. We want to provide this tool to help motivate the average person become more politically involved and to express themselves using different social media and find other people interested in politics.

1.2.1 Question Formulation

The question this master thesis attempts to find a solution for is: **"How to create a real-time feedback system for the political environment with**

automated trend-analysis?". The application will detect and gather written posts from social media Twitter concerning Swedish politics. The written text will be analyzed using Saplo's algorithms to be able to create a trend for the political party or leader in question.

We are hoping this type of system will make the average person more involved and interested in politics by making their opinions visible on the application and getting instant feedback. We are also hoping that this application will motivate politicians to become more active at using social media.

1.3 Purpose

As mentioned in "Goals" we want to provide a tool to make the average person heard in the political environment as well as using their opinions to visualize trends for the political parties. There is also the business aspect of developing this type of application as Saplo wants this application for marketing purposes. Saplo is very proud to have developed algorithms that could analyse large quantities of data and are therefore eager to show the world their capabilities.

1.4 Delimitation

This master thesis will only focus on the social media Twitter due to time constraints. The application will be developed in such a way that it can be extended to other social media, news sites as well as other information media even though it will not be implement during this project.

The application developed during this project will solely cover the development and implementation of the visualization of political trends as they are presented on Twitter. The text of each tweet will be sent to Saplo's API for sentiment analysis which Saplo will provide and any adjustments on this part will be handled by them. Therefore the quality of the analysis is something that we can't affect.

Chapter 2

Method

The project will be divided into four different phases. First, due to the critical time aspect of the master thesis it is important to have a structured time plan in order to be able to execute the project in an efficient manner. The time plan will be constructed by using a work breakdown structure (WBS) and a Gantt chart. The WBS gives a more detailed overview of what parts the project consists of and the Gantt chart organize the project parts according to their importance and order in which they need to be executed [14].

In the second step we need to research what type of programing languages that are suitable for this project. There is always a balance of using the most modern technology and how many people that actually are able to use it because of for example out dated web browsers. The research phase will continue with a quick competitor analysis and a target-group analysis, to get an idea that other companies offers. It will hopefully make it easier for Saplo to decide how to market the application and at whom it is aimed towards. This will play an important role in creating an interesting interface appropriate for that specific group of people.

After having specified the requirements for the site, both functional and non-functional, the design phase will begin. What we know today is that Saplo request the following components on the site: first, a line chart to plot the trend index that Saplo provides. Secondly a stream of tweets needs to be present on the site to give a real-time impression. Saplo emphasized the importance of real-time due to marketing purposes, they want to display how fast they are able to analyze the information from Twitter. This requirement limits the design slightly but otherwise they were very open for ideas. The first stage will be created with pen and paper in a quick and dirty manner, just to get an idea what components that are essential for the design. The design will further evolve by user-evaluations and testing.

The implementation phase will be performed in parallel to both some minor design choices and testning of the application. Testing and validation of code will be performed in parallel to the implementation, by our own experience this method has proven many times to be the most effective and delivers overall better and more reliable results than projects where testing is left to the very end.

Chapter 3

Planning

The critical time aspect of this master thesis required a well structured and defined time plan together with a risk analysis. The first step in planing is to have a clear and defined view of the task. We choose to use the work breakdown structure (WBS) approach followed up with a Gantt chart.

3.1 Work Breakdown Structure

Work breakdown structure, or WBS for short, is a useful method to break down a project into smaller and more easily supervised activities [19]. The project was first broken down in four major parts: pre-study, design, implementation and testing. All these parts were then broken down over and over until a clear picture occurred with all the parts the project consisted of. Without going into to much detail, the following groups were created:

Prestudy

- Competitor analysis
- Target group analysis
- Literature study
- Design and concepts

Implementation and Further Design

- Prototype
- Re-design
- Implementation and testing

3.2 Gantt Chart

The WBS technique gave a clear view on what parts the projected consisted of but not how much time each activity would take. To get a better understanding, a Gantt chart was used. Gantt charts sorts all activities depending on their start date and their importance [20]. The Gantt chart gives projects a good structure and overview.

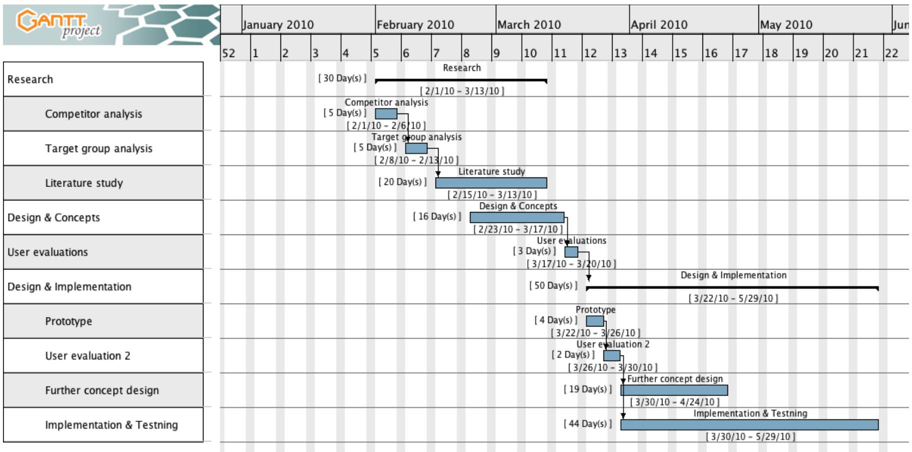


Figure 3.1. Gantt chart

Chapter 4

Concept

By using the information gathered from Saplo as a foundation for the application it was time to start researching to find a good concept for the application. The first step in the search for concept is a competitor analysis for create a clearer picture of how to make our vision unique and exciting. Later, the brainstorm method will be used to find possible designs of easy and informative interfaces which will meet Saplo's requirements.

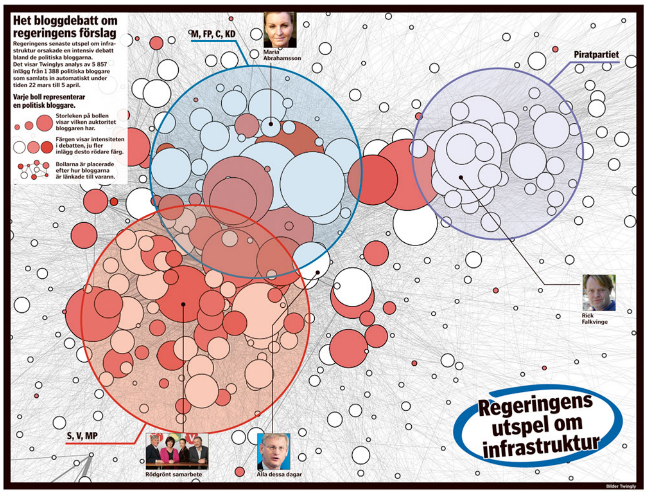
4.1 Competitor Analysis

When developing new types of services or products it is always a good idea to be aware of the competition. The estimates are that the application would be quite unique due to Saplo's highly advanced text analysis algorithms but never the less it is important to keep oneself updated on what the competition can offer in order to keep an competitive advantage.

The application is very unique, analysing social media for the Swedish election. Due to this very niched nature of the application it was difficult to find any competitors that were doing the exact or similar thing. There are however some competitors that have similarities to what this master thesis hopes to achieve.

4.1.1 Twingly

Twingly is a search engine for blogs. Their main product is called "Twingly Blogstream" and is created to increase the SEO results on the sites that uses that product [23]. This is done by linking to back to different blogs that links to the relevant site in question, these blogs will get more attention and increasing their values and therefore the SEO on the site in question [24]. Twingly has recently start using their search-engine technology analyzing the Swedish election [2] as can be seen in figure 4.1. However, the analysis is not semantic but purely a visual presentation on how blogs are connected.



Het bloggdebatt om regeringens förslag

Regeringens senaste utspel om infrastruktur orsakade en intensiv debatt bland de politiska bloggarna. Det visar Twinglys analys av 5 857 inlägg från 1 388 politiska bloggare som samlats in automatiskt under tiden 22 mars till 5 april.

Varje boll representerar en politisk bloggare.

Storleken på bollen visar vilken auktoritet bloggaren har.

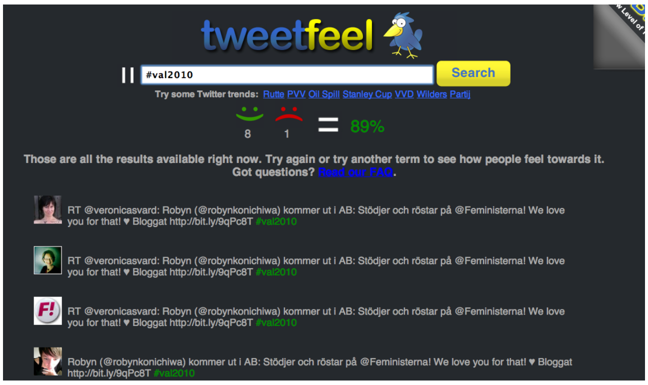
Färgen visar intensiteten i debatten, ju fler inlägg desto rödare färg.

Bollarna är placerade efter hur bloggarna är länkade till varann.

Figure 4.1. Twingly

4.1.2 Tweetfeel

Tweetfeel markets themselves as "Real-time Twitter search with feelings using insanely complex sentiment analysis." [29]. They let the user enter a search word and the search Twitter and Tweetfeel's own database to find relevant tweets and analyzes them. For the best result they suggest that the user searches for popular movies, celebrities, TV shows and companies but the user can choose to search for whatever he or she desires.



tweetfeel
#val2010
Search
Try some Twitter trends: Rutte PVV Oil Spill Stanley Cup VVD Wilders Partij
8
1
= 89%

Those are all the results available right now. Try again or try another term to see how people feel towards it.
Got questions? Read our FAQ.

RT @veronicasvard: Robyn (@robynkonichiwa) kommer ut i AB: Stödjer och röstar på @Feministerna! We love you for that! ♥ Bloggat http://bit.ly/9qPc8T #val2010

RT @veronicasvard: Robyn (@robynkonichiwa) kommer ut i AB: Stödjer och röstar på @Feministerna! We love you for that! ♥ Bloggat http://bit.ly/9qPc8T #val2010

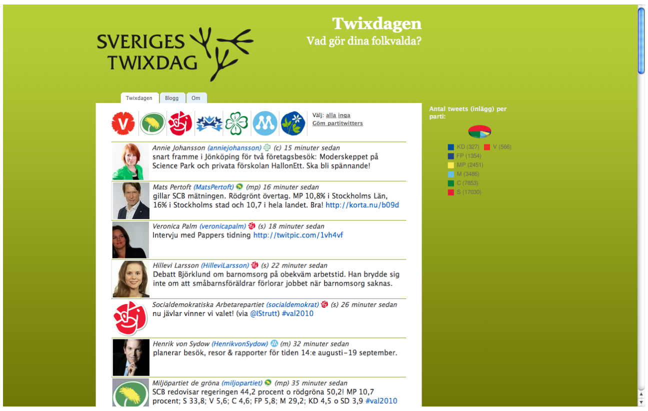
RT @veronicasvard: Robyn (@robynkonichiwa) kommer ut i AB: Stödjer och röstar på @Feministerna! We love you for that! ♥ Bloggat http://bit.ly/9qPc8T #val2010

Robyn (@robynkonichiwa) kommer ut i AB: Stödjer och röstar på @Feministerna! We love you for that! ♥ Bloggat http://bit.ly/9qPc8T #val2010

Figure 4.2. Tweetfeel

4.1.3 Twixdagen

Twixdagen is created by a politically independent magazine, they want to contribute to political coverage that consent to dialog and openness without renounce the review [11]. They created a tool that gathers tweets posted by political parties as well as their members and displays basic statistics on how many tweets every political party has. In contrast to our application Twixdagen has no underlaying system that analyzes the incoming messages from Twitter nor does it take in tweets from average users.



Annie Johansson (anniejohansson) (c) 15 minuter sedan
snart framme i Jönköping för två företagsbesök: Moderskeppet på Science Park och privata förskolan HallonEtt. Ska bli spännande!

Mats Pertoft (MatsPertoft) (mp) 16 minuter sedan
gillar SCB mätningen. Rödgrönt övertag. MP 10,8% i Stockholms Län, 16% i Stockholms stad och 10,7 i hela landet. Bra! http://korta.nu/b09d

Veronica Palm (veronicapalm) (s) 18 minuter sedan
Intervju med Pappers tidning http://twitpic.com/1vh4vf

Hillevi Larsson (HilleviLarsson) (s) 22 minuter sedan
Debatt Björklund om barnomsorg på obekväm arbetstid. Han brydde sig inte om att småbarnsföräldrar förlorar jobbet när barnomsorg saknas.

Socialdemokratiska Arbetarepartiet (socialdemokrat) (s) 26 minuter sedan
nu jävlar vinner vi valet! (via @lStrutt) #val2010

Henrik von Sydow (HenrikvonSydow) (m) 32 minuter sedan
planerar besök, resor & rapporter för tiden 14:e augusti–19 september.

Miljöpartiet de gröna (miljopartiet) (mp) 35 minuter sedan
SCB redovisar regeringen 44,2 procent o rödgröna 50,2! MP 10,7 procent; S 33,8; V 5,6; C 4,6; FP 5,8; M 29,2; KD 4,5 o SD 3,9 #val2010

Figure 4.3. Twixdagen

4.1.4 Aitellu

Aitellu's idea is very similar to Saplo's, they both focus on automated text analysis [3]. Aitellu has previously provided the viewers of the Swedish song contest with information on that the blogs thought about the different artists. Aitellu's production manager, Anders Bjurström, says that their technology will be used during the election as well [2].

4.1.5 Conclusion of Competitor Analysis

There are a few similar alternatives to what we hope to develop, but none of them are exactly the same. Twingly and Twixdagen both gathers and filters information that makes it much easier for the user find relevant information, but they are not using any intelligent algorithms to analyze the text. The alternatives that on the other hand uses analytical algorithms are Tweetfeel and Aitellu. They can both analyze text and give an understanding if they are positive or negative. Tweetfeel does not target the Swedish election nor any other specific subject. Aitellu is impossible for us to truly analyze because they haven't launched any system for the Swedish election and still it is not sure that they will.

The competitor analysis helped us realize what makes our vision unique. Our vision both filters the information for relevance to the Swedish election as well as analyzing it in real-time. In the same instance as a tweet is detected on Twitter, Saplo will analyze the content of that tweet and send the tweet and the result of the analysis directly to our application. This will hopefully make our application very interesting during for instance debates on TV where people's opinions can be viewed as soon as they are written. This could even help the politicians that are live om TV by sorting the tweets by negative or positive they can see what people believe is good and bad. For these reasons we believe that our vision can stand up to the competition.

4.2 Target Group Analysis

To get a better understanding of our target group we performed the method "Extreme Personas". The method is created by a company named IDEO [12], they made themselves famous for their creative ideas and methods for coming up with them. The idea is to create a few very different fictional persons that would use the application in very different ways.

4.2.1 Extreme Personas

The method extreme personas was used in order to get a better understanding of what type of target group to focus the application at. Four different extreme personas were created, each with very different needs and views on this type of tool. To shortly summarize the output of the method, the following four persons were created:

- Anna the journalist
- Jessica the twelve year old schoolgirl
- Tomas the event manager
- Maria the senior citizen

These are all persons that would use this type of application in very different way. The following describes the different personas that the study resulted in:

Anna is very interested in gathering statistics for her newspaper. She is a person on the run and wants get the latest numbers presented in an quick and easy way for her to see. For Anna the option to get the latest trends directly to her phone is very important when she is out on her journalistic missions.

Jessica is a young schoolgirl that has just started getting curious about politics and have therefor started searching for means to learn more. Jessica is so far rather disappointed, every source of information about politics is very boring and written with difficult words that in her opinion only adults could understand. She cant understand why everything needs to be so difficult with strange statistics, numbers and weird words. She want to find a place where she can take part of Swedish politics and in a more playful way learn more what everything is about.

Tomas works as a events manager and has this year been asked to organize an event for the election night watch. Tomas finds it important to have a product that can be put on large projectors and give the participants of the events real-time updates about the election. It is important for Tomas that the information fits on a screen while at the same time is easy for the viewer to read and understand what's happening.

Maria has been interested in politics all her long life but is not very familiar with modern technologies like the Internet. She therefore is similar to Jessica in that she wants it very basic and simple without to much advanced functionality and complicated interfaces. However, in contrast from Jessica, Maria still wants detailed information about her favorite political party and how they are doing. The information she would want is a collection of information in one place, like news and SIFO surveys. This would be very useful for a senior citizen whom is not used to the Internet, since they could get a lot of information in just one place. Maria also likes clear buttons with colorful icons because it is easier for her eyes to read.

Saplo had very specific demands, they wanted their application to be used by everyone. The personoas helped in showing Saplo how very different ways people could use this type of application. After discussion with Saplo they still wanted to accommodate everyone and not prioritize a specific group of people. This means that Saplo wants the application to be able to be displayed on a high-definition TV screen. This will play an important aspect for our design.

Saplo's desire to accommodate everyone could mean problems in finding a design that suits everyone. The design therefore needs to compromise between all of them and at the same time trying to keep it professional due to the political nature of the project.

4.3 Alternative Views

Saplo gave some guidelines for how the analysis tool could look or at least some components they wanted to include. We however wanted to explore other possibilities for the application or at least suggestions for future development. These suggestions could also work in parallel to a more "professional" site as well, for instance as an entry point to the page. The following two concepts are a result of several brainstorming sessions.

4.3.1 SapGarden

SapGarden is basically a bar chart in disguise, a very visual presentation of the Swedish election. It consists of a few different elements; the cloud represents twitter or the Internet cloud and the raindrops are tweets. The political parties are represented by flowers. Alternatively the party alliances are the flowers and the individual parties make up the roots to the flowers. In the later option both the trend of the individual parties and their alliances would be clearly visualized.

The gardener is holding a measuring rod in his hand with "100%" written on it. Here the user can see how the trends are distribute among the different parties or alliances.

Whenever a user on Twitter writes an opinion concerning the Swedish election their tweet will be analyzed by Saplo's algorithm and presented in real-time as a raindrop. That raindrop will be animated to fall down over the relevant party/alliance (that is a flower) as to simulate rain. If the tweet is positive it will help the flower grow taller. The raindrops could also change color depending on if they are positive, negative or neutral. The user of SapGarden could hover the mouse over the raindrops and thereby slowing down the decent of the raindrop and present the text of the tweet allowing the user to read it.

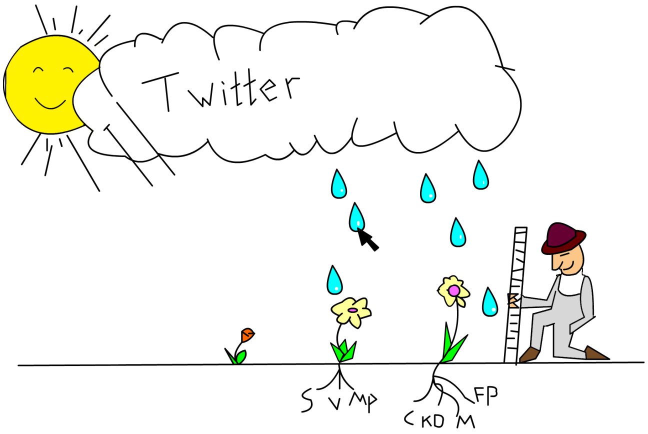


Figure 4.4. SapGarden

4.3.2 TweetBird Express

TweetBird express consists of three elements and is an attempt to visually emphasize that Saplo's algorithms work in real-time. As soon as a person posts a tweet regarding the election on twitter, a bird will enter the screen flying in from the left holding a scroll which will have the party's logo on it which will represent the tweet. The bird will drop the scroll into the "Saplo-machine" (the orange box attached with balloons) where the text will be processed. The message (tweet) is then presented with text together with previous tweets in the left column. To the right will be a line chart displaying the parties trend index over time. The relevant political party's trend will at the same time be affected as the tweet is presented to the left so that the user realizes the connection between them.

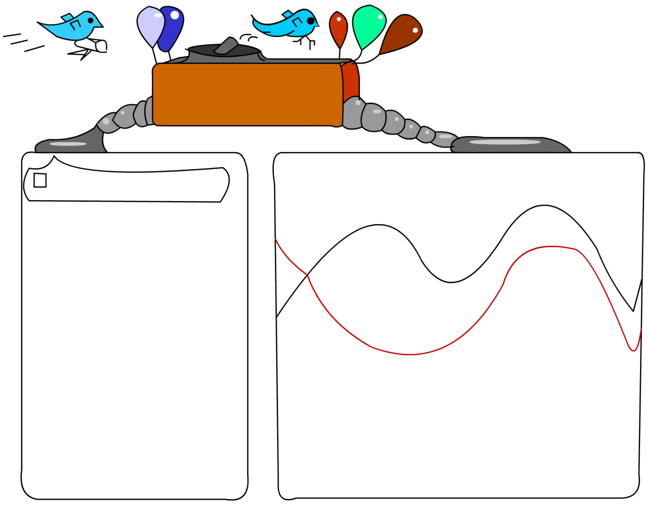


Figure 4.5. TweetBird express

4.4 Main Application Concept

Both SapGarden and TweetBird express are concepts for alternative views, however the core application is aiming for a more professional look.

Some restrictions were put on the design given the wishes from Saplo. They wanted to be able to use the application both on a normal computer screen but also as a something which could be put up on a high-definition TV in for example a convention centre and still be visually pleasing and informative. The big-screen television, even though being physically large, will have a lower resolution than many computer screens and since it will be in a larger setting it will render it impossible to scroll around on the page. Therefore we had the challenge of presenting a lot of information in a limited space.

4.4.1 Concept 1: Trend Visualisation With Map

Keeping it as simple as possible with only a line chart and the messages below would give a application a very minimalistic look and feel. A map was added so that people could observe where in Sweden the tweet was posted. The different regions could also change political color on the map, depending on in what positive political direction the tweeters are writing in that region.

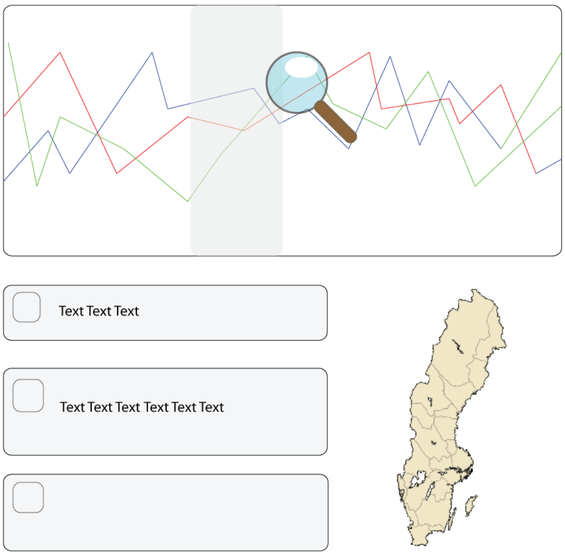


Figure 4.6. Concept 1: Trend Visualisation With Map

4.4.2 Concept 2: Dropping Tweet-Spheres

An extended version of our first mock-up was the dropping tweet-spheres. Whenever a tweet is detected by the application it will be represented by a sphere. That sphere will place itself in the chart and also drop a copy of itself down to a slanting "ramp" below. The sphere will roll down to the end of the ramp and then drop down and form a message. This will be a fun and intuitive way for the user to get a visually representation of what’s happening in real-time. The user could also hover over the spheres to read the tweets that affected that point in time.

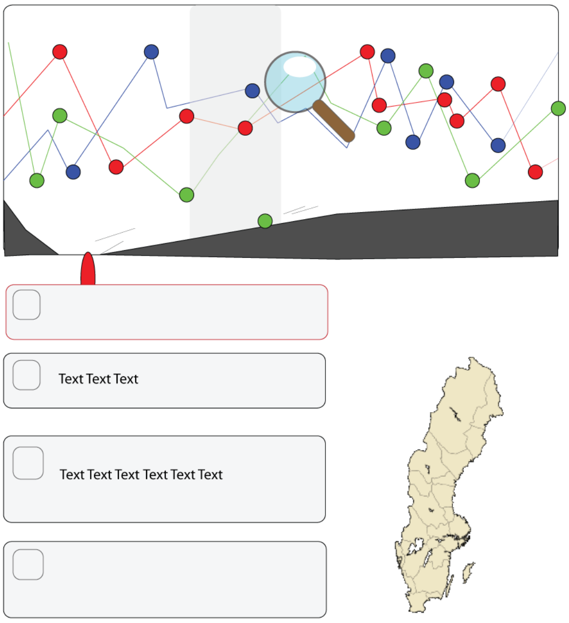


Figure 4.7. Concept 2: Dropping Tweet-Spheres

4.4.3 Concept 3: Sphere Attracting Chart

This concept is also an extension of the first concept. Here each new tweet will be ploted in a special area on the line chart's right hand side. After a fixed set of time all the spheres in this area will be "sucked" together with the ones from their corresponding party forming a larger ball. The size and position of this new ball would depend the number of tweets which it consists of and their average sentiment respectively. The ball would then be placed in the "normal" area of the line chart and the process would start over.

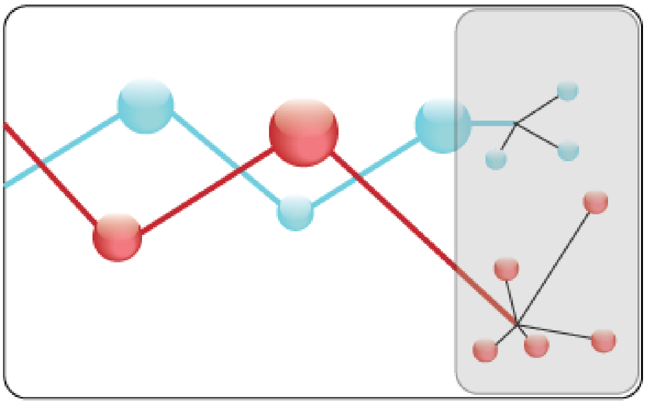


Figure 4.8. Concept 3: Sphere Attracting Chart

4.4.4 Concept 4: With Information Box

The fourth concept also consists of three different elements; the tweet-flow to the left, the line chart and a information-box down to the right. The tweet-flow and the line chart has the same function as described above in section 4.4.1 with the only difference that the line chart has been made a bit narrower in order to give more space to the tweet-flow on the left which is now able to reach higher. The new element is the information-box which is placed beneath the line chart. This area will be interactive in the way that it will have two different layouts. The first layout, which will be initially displayed when entering the page, will have a table showing how many tweets and the current index for both the political parties and their party leaders. In addition to this view there will be one for each party which can be accessed through a tab system. The party tabs will consist of among other things a pie chart displaying the distribution of positive-, negative- and neutral tweets for the party.

This layout contains a lot of information with a small area and would be very appropriate for a TV-screen.

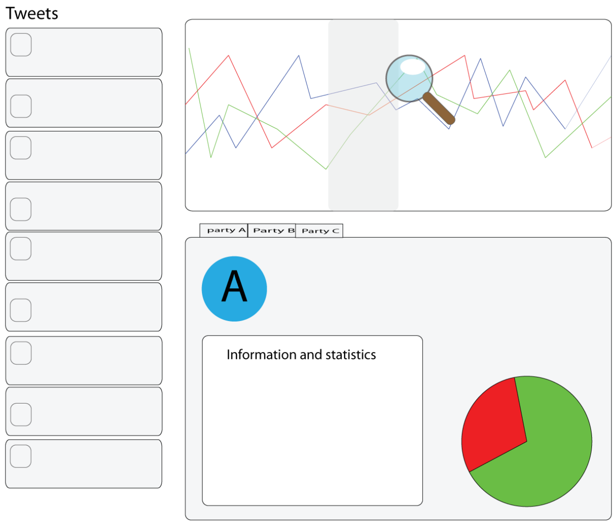


Figure 4.9. Concept 4: With Information Box

4.5 User Evaluation

After developing the first concepts it was time to test them out in a user evaluation to make sure people understood them. This evaluation was preformed with a method similar to paper prototyping; instead of creating the mockups on normal pieces of paper they were designed in Adobe Illustrator. The test-subject observed our concepts on a computer screen and pointed out the different elements. They first gave us input on what they thought the different elements were supposed to do and later we explained them to give the test subject a better understanding in order to give better input.

4.5.1 SapGarden and TweetBird Express

All the test subjects experienced both SapGarden and TweetBird express to be very innovative and smart. But they weren't as self-explanatory as we first hoped. Mostly people didn't realize why the raindrops were falling, why the birds came flying with messages nor what really triggered that event. After the designed was explained people found the design to be a fun way to illustrate the process of first detecting the tweet, analyse it and display it to the user.

4.5.2 Main Concepts

To summarize the user-study from the main concept evaluation, people generally liked first and last of our concepts. The first concept with its clean, minimalistic look only has the line chart, tweet flow and a map of Sweden. This concept was much appreciated due to its simplicity. All three sphere charts were however considered more confusing than visually helpful, which was rather disappointing. The last concept, number four, was considered to be very informative but some test subjects expressed concern that it would contain too much information if not designed correctly.

4.6 The "Final" Concept

Based on our target group analysis, competitor analysis, concepts and evaluation the "final" concept started to take shape. During our user-study we found out that the map of Sweden wouldn't work very well. This feature requires that the tweet contains geographical information, known as geo-tagging. This is a feature which Twitter supports but is however turned of as a default and must be actively enabled by the user. Today approximately only 0.23% of all twitter-users have enabled the function [4]. It was deemed to give a skewed representation of the political views dispersed over a map of Sweden.

One alternative know was to just remove the map of Sweden in our version one and make it even more minimalistic. But after reconsidering Saplo's request from the target group analysis we wanted to put more information on the first page to be suitable for a TV-screen. This lead us to concept number four. This concept contains just like concept one both the tweet-flow and line-chart but have also added a information box. The users that tested our designs thought this could work very well, even though some people were worried that it could get confusing with to much information if not designed correctly.

Concept four will be used as a foundation for further development and base for our application. The concept probably will need more work and research before establishing the concept for the application, especially the information box need to be evaluated.

Chapter 5

Research

The research was divided into a few sub-tasks; competitor analysis, target group analysis, literature study and user evaluations. Both target group analysis and competitor study has already been described in the section "Concept". The following describes the research needed in finding the best technical means to complete our task.

5.1 Choosing a Platform

Before starting the research phase a platform needed to be decided. When choosing a platform, the nature of the project is very important to take into consideration. The analysis tool should both be promotion for the company and a source for information considering the Swedish election. In both aspects the importance to reach as many people as possible is highly essential. Even though mobile platforms are a hot topic and their popularity is growing like the iPhone OS and Android the decision was made not to start developing specifically for these types of platforms. Instead a web application seemed more appropriate due to its accessibility. However, the mobile platforms are still very important and they will be taken into consideration when developing the site. Research needs to be performed on what different programing languages the major mobile devices can handle.

5.2 Twitter

Twitter is a social network containing of mini-blogs or as Twitter describes it: Twitter is a real-time information network powered by people all around the world that lets you share and discover what's happening now. Twitter asks "what's happening" and makes the answer spread across the globe to millions, immediately." [25]

There are limitations that needs to be taking into consideration when working against Twitter. First of all, the tweets are limited to 140 characters which Saplo fears can cause a lower accuracy in the analysis. Another important issue working against Twitter was how to gather the information in large quantities due to specific keywords. There are two different options: FireHose and the Streaming API.

5.2.1 FireHose

Fire-hose gives full access and unlimited access to all information on Twitter, around 50 million tweets per day. Previously it was only the giant search-engines like Google that got full access to Twitter. But in December 2009 Twitter announced that they would open fire-hose to more and more developers and start-ups companies [18]. However, its still difficult to get access to this service so it wont be relevant for this particular project.

5.2.2 Streaming API

There are currently over 50.000 applications using the Twitter streaming API [18], this amount of applications have lead to some limitations from Twitter. The streaming API does not give the same full access as the fire-hose. However, the streaming API gives developers near real-time access to sub-sets of tweets . This service is very suitable for our purposes concerning the Swedish election due to the use of specific keywords.

5.2.3 Collection Framework

In order to collect tweets Twitter has created a streaming API which operates purely over HTTP. Several different frameworks, in multiple programming languages, have been developed by independent users to simplify the process of setting up a connection to the API. These vary greatly in both implemented functionality and documentation.

Saplo requested that Java should be used as they already use it in several other applications. The available frameworks were researched independently by us and Saplo and both decided on Twitter4J due to positive reviews and the relative ease of getting it up and running.

5.3 Front-End

Saplo tries to market themselves as a young and innovative company with new and fresh ideas. It's therefore important trying to match their image with the site's design while still keeping the site available for as many people as possible. These factors are very important to take into consideration when choosing a programing language for the site. It should both be new, fresh and yet functional with good performance. To meet these requirements, the following program-languages are commonly used for web development and will be evaluated with pros and cons to see how suitable they are for our purposes.

5.3.1 Flash and ActionScript

Flash is developed by Adobe and is often used to develop so called Rich Internet Applications (RIA), these applications share many characteristics with desktop applications. Flash can for example create visual appealing websites, games and commercial widgets. Flash is strongly connected to the script-based language ActionScript. A language that was initially designed to controlled vectors in Abobe Flash but has evolved to a more sophisticate language capable enough for games and RIA. Adobe presents flash: "Adobe Flash Player is a cross-platform browser-based application runtime that delivers uncompromised viewing of expressive applications, content, and videos across screens and browsers. Flash Player delivers breakthrough web experiences to over 98% of Internet users." [1].

5.3.2 Silverlight

Silverlight is Microsoft's answer to Adobe flash, used for web, desktop and mobile applications. It is powered by the .NET framework and is supported by most of the modern web browsers [5]. However, silverlight is not yet standard in all web browsers and therefore requires the user to download and install a plug-in.

5.3.3 JavaScript

Javascript is an well established object-oriented scripting language for the Web and shouldn't be confused with Java for desktop application. It is supported by all major web browsers and used on millions of web sites though out the web [14]. JavaScript is lightweight and executes directly in the browser without any preliminary compilation.

5.3.4 HTML5

As the name indicates HTML5 is the fifth version of HTML and will replace the present versions (HTML 4.01 and XHTML 1.0) as the standard for the web. HTML5's built in functionality that works in similar way as flash have led to a discussion whether or not HTML5 will lead to the end of flash [6].

5.3.5 Front-end Summary

In order for the site to be accessible for as many people as possible both Silverlight and HTML5 didn't seem like the way to go. They have both great potential for the future, especially HTML5 as the new standard for the Web. However, it is still a young technology and many older web browsers doesn't support them. Flash on the other hand seemed like an very interesting choice. Its possibility to design very interesting and innovative interfaces would give us great potential to design something very interesting and informative for Saplo that would fit their image. However, the iPhone has no support for flash. Even though we are not developing or optimizing the product for mobile platforms specifically we don't want to exclude them by using a language that they can't handle. After having excluded these three options we were left with Javascript which seemed like a great choice for our purpose.

5.4 Back-End

The back-end will consist of a Java client that uses the Twitter streaming API to setup a connection to Twitter. A filter with specific keywords relevant to the Swedish election was applied to the stream in order to only collect relevant tweets. The client takes the tweets and stores them into a MySQL database while at the same time sending the tweet text to be analyzed by their algorithms.

Saplo has previously worked with similar back-end solutions and will therefore set up the major part of the system. The back-end system will also consist of an Apache web server that will handle the information between Twitter, Saplo API and the website.

5.4.1 Webserver

An Apache web server has two functions; The first is to serve as a normal web server and process the PHP scripts as well as creating JSON-files which are continuously requested by each client in order to have the latest tweets.

5.4.2 Database

A database also needed to be setup in order to store the collected tweets and statistics connected to this. For this MySQL was chosen as it is free under the GNU General Public Licence and also in use by Saplo in other projects.

5.4.3 Webscripting Language

As our page needs to be dynamic and updated continuously it is not enough to just use HTML/XHTML, therefore it was important to find a scripting language appropriate for our needs. The two languages which were considered were PHP and ASP .NET

PHP

PHP is a general purpose scripting language which was designed for web development and its code can therefore be integrated directly into HTML code [16]. It can be run on most web servers and can be connected to any relational database management system (RDBMS). All PHP code is parsed in realtime when the code is requested. Several frameworks such as CakePHP and Zend Framework have been developed in order to structure the development of applications by implementing a model view control (MVC) architecture as well as security, caching and form validation features. PHP is both open source and free to use.

ASP. NET

The web development framework ASP. NET is developed by Microsoft and is a successor to their ASP framework. When a ASP .NET page is requested for the first time it is compiled at runtime and then cached in order to speed up subsequent requests [26]. It has built in support for both authentication, account and role management. It is also very scalable [26].

Back-End Summary

Both languages meet the requirements of this project without any problems but the decision was made to use PHP as the development language. The two main reasons for this was the open source aspect of PHP and a greater familiarity with the language from previous projects.

There also seems to be the general consensus on the Internet that PHP has better support for the integration of the database management system MySQL which was also used in the project and therefore was a plus for PHP.

5.5 Research of JavaScript Frameworks

Since JavaScript is integrated into all web browsers and offers developers the possibility to create dynamic websites with enhanced user interfaces it is now used on 99.46% of all websites [22]. To facilitate the project as much as possible a research study began to investigate if there existed any frameworks to help us with the project. The research continued with the search for a line-chart framework, as mentioned previously Saplo requested an interactive line-chart to display the trends in real-time. A secondary objective was to also find a framework which could handle static charts.

5.5.1 Selecting a JavaScript Framework

JQuery is a free open sourced Javascript library which is designed to simplify HTML scripting on the client side. It allows for DOM traversal, CSS manipulation, event triggers, effects and animations as well as handling AJAX requests. Due to its lightweight nature, speed and cross platform capability it has become the most popular Javascript Framework and is used by over 31% of the 10,000 most visited websites [22].

It was decided to go with JQuery as the overall JavaScript framework. Even though there are others frameworks like MooTools, Prototype and Script.Aculo.Us which can perform most of the same things as JQuery and in some cases are compatible with older web browsers. The popularity, community and the ability to develop plug-ins weighed heavier for JQuery. But the main reason for our decision was however that we had previous experience with JQuery and knows that it will for fill our purposes.

5.5.2 Chart Frameworks

Plotkit

Plotkit describes themselves as "Chart and Graph Plotting Library for Javascript. It has support for HTML Canvas and also SVG via Adobe SVG Viewer and native browser support." [17]. Flotkit has support for many different types of graphs and charts, like: line charts, bar charts and pie charts. The homepage contains a lot of different examples, tutorials and an active community to help developers get started using Plotkit.

Flot

Flot is a graph library entirely develop in Javascript and jQuery (more about jQuery later). The Flot plug-in works with all major browsers and uses the HTML canvas tag and the excanvas plug-in for Internet Explorer [15]. It is open sourced and has evolved with the help of an active community. In addition to its community it also has a well documented API and the possibility for users to create their own plug-ins. It is limited to displaying line- and bar charts making it necessary to find an alternative solution for pie charts. The interactive and customization options also suit our needs in that it is possible to display multiple charts on one page as well as multiple series in each charts which can be turned on and off.

Emprise JavaScript Chart

Is a javascript framework which allows for creation of both line- and bar charts which can be made interactive while at the same time being visually pleasing. They are also compatible with all major browsers. Ejschart offers a lot of different functionality such as interactive charts, mouse events, axis scaling, auto zooming and supports XML-formatted data and loads data directly on the site [7].

Google visualization API/Google Charts

Google offers a collection of different charts, both interactive and static [10]. They are designed to be extremely simple to use. An example of this is that for the static charts Google creates a PNG image by a HTML request with certain parameters. A very simple way to implement charts on a web-site. The interactive charts are created by using SVG alternatively Flash.

Summary Chart Frameworks

Both Ejscharts and Google charts didn't for fill our requirements for the main line chart on the site. Ejschart looks very appealing and is written in pure JavaScript and compatible with all major browsers, which is exactly what we are searching for. It was however ruled out as an option due to the fact that the 'free' version had several restrictions such as a watermark on it as well as there only could be 2 series per chart which isn't enough for our purposes. Google charts was ruled out quickly due to that google uses flash and SVG for their interactive charts, these techniques are not an option for our site. The static Google charts on the other hand could come in handy due to their simplicity and that they are constructed by a HTML request and displayed as a PNG file.

Deciding between Plotkit and Flot was on the other hand very difficult. Both frameworks seemed to for fill our requirements. Therefore both visual styles and documentation was compared to settle the differences. Flot was chosen as the framework to create the line-chart on the site, mostly because of the pedagogical examples and the overall impression of the framework.

5.5.3 Data Transfer

In order to update the page in a effective matter a variety of techniques is needed. As mentioned above PHP is used in order to make the page dynamic. However, PHP only get the latest info when the page is loaded and the site requires to be continuously updated. The site therefore needs to use some kind of data transfer protocol. The different methods needed are described below. They are all chosen after dialogs with Saplo in order to get the back-end and front-end to communicate in the best manner possible.

Ajax

Ajax is short for asynchronous JavaScript and XML. Ajax makes it possible for websites to asynchronous communicate and receive data from the server without interfering with the display of the web page [9].

Cross-Domain

Due to security risks which can arise when requesting information from different origins JavaScript has implemented a Same Origin Policy [28] . This limits the Ajax requests to be from the exact same domain which in our case doesn't work since the webpage and the updated tweets come from different servers. In order to circumvent this problem an application called Ajax-Cross-Domain was used [27] . This allows the website client to make the Ajax request to the same domain as the webpage is on and that script then fetches the requested data by forwarding the request to the desired site.

JSON

JavaScript Object Notation or JSON as the acronym stands for is an "data-interchange format" and is most commonly used for serializing and transmitting data over networks in web applications. It is purely text-based and can therefore be naturally read making it easy for humans to read and write [13] . As the name implies it comes from a subset of the JavaScript language and uses the same structure in representing data structures and associative arrays.

Chapter 6

User Testing

All in all, there were three major user testing periods of the project: the pre-study which took place after our first concept design, one during our implementation and one in the end phase of the project.

The first user-evaluation was performed after the first concepts were completed, this study is briefly described under the section "Concepts". Different layouts were presented for the test-subject and they explained how they experienced the design. That user-evaluation help us understand how to place our elements in a pedagogical matter and was a foundation for the further concept development (see next section).

The other two user-studies was designed to letting our test subject perform certain use-cases. They the use-cases designed to cover all the critical functionality on the site. However, the first step in these user-tests where to just let the user play around on the site and think out loud.

By just being an observer with pen and paper we were able to collect valuable information on how users interpret the site and its functions. After just letting the users play around on the site for a while we gave them orders to complete certain task or use-cases as mentioned earlier. The user when navigated around trying to complete the required task. It was noted how different people solved the "problem" and how they were thinking.

An example of these type of use-cases could be to compare only two parties in the line chart and explain the information that the graph presents. Here again the user thought out loud and all their thoughts was put down on paper.

Chapter 7

Further Concept Development

As described in the Concept section an initial concept was developed consisting of the three parts: twitter flow, a line chart and an information box. These three parts however needed to be refined in their design and the information box also needed to be further developed.

This further development was done in parallel with the implementation phase and user testing phase, which will be discussed further below. By working on these three phases of the development in parallel we were able to get better and more accurate feedback from the test subjects as opposed to testing and finalizing the design based on paper prototypes.

7.1 The Twitter Flow

The concept of the twitter flow was fairly clear from the beginning, this in large part to its simplicity as well as the fact that we wanted to keep it fairly similar to that of Twitters on layout in order to keep the recognition level for the user.

Just like on Twitter each post will contain the user image, name, message as well as a reply link which the user can use in case the user wants to send a reply to a post. It will also contain when it was posted in contrast to Twitter where it says how long ago it was posted. The reason for this change was that our application will update continuously as opposed to twitter which recalculates the lapsed time when the page is refreshed.

In addition to these "standard" features the logo of the parties mentioned in the tweet will be displayed in the top right corner next to a thumbs-up/thumbs-down figure which will indicate the sentiment of the tweet. The hands are also color coded, so that a thumbs up is green and thumbs down is red, so that the user can get an even faster impression of the tweet's sentiment.

Clicking on the reply link will open up a new Twitter window which, if logged in, will be centered on the dialog box with information to send a reply to the given tweet.

7.2 The Line Chart

The line chart was the second thing which Saplo requested from near start. This also seemed fairly straight forward at a first glance but required a lot more consideration.

The first obstacle we met was the number of series which should simultaneously be presented in the graph. As mentioned earlier the application is going to track 11 different parties, the seven in currently in the parliament and the four largest challenging parties. In addition to these 11 parties we also wanted to be able to show the trends for their party leaders which is another 12, since Miljopartiet has two, making a total of 23 different series. To display all 23 series at the same time would however make it impossible to distinguish the series from one another. Therefore its important for us to keep the number of active lines down to a reasonable amount in order to keep the readability high. This could lead to an restriction on how many active lines the user can show at the same time.

This however caused a dilemma as to which series to display as the page is supposed to be politically independent. After a lot of deliberation it was decided to use seven parties which are currently part of the parliament as the default and give the user the option to manually select the series which they desire to view.

Just above the line chart a options button was placed which when clicked will present an options area. This will contain a list of all the available parties and their leaders alongside with check-boxes. The parties and their leaders were divided into two separate groups but were placed in the same order as to facilitate the localization of the party leaders. In addition the abbreviation of for the respective party was places after each party leader in order to make it even clearer.

When a check-box is selected/deselected the change is made immediately in the the line chart giving the user instant feedback. A limit to the maximum number of simultaneously displayed series was set to seven due to the limit of space in the charts legend. In the case that a user tries to exceed this limit a warning dialog is displayed informing the user of the limit.

In addition to the series and their check-boxes two buttons were added: one to deselect all series and thereby making it easier for the user to then select only the ones which they are interested in and thereby removing the excess steps of deselecting them one by one. The other button selects all the parties making up the parliament and deselecting the others and in effect resetting the chart to its original status.

In the top left corner a chart legend was added displaying which elements are selected and their respective colors. In order to compensate for the fact that we exceeded the guidelines to have a maximum of five series in a single chart the legend was made to be interactive in aspects. The first thing is that when the user overs over the line chart a vertical red line is drawn over the chart indicating the given position in time space and the given value is shown in the legend to the right of each series name. This way the user can quickly and easily retrieve the precise value for a series at any given time in the chart.

The second thing is that each series name is clickable in the legend. When a series is clicked on two things happen depending on the current state. If it's the first time a user clicks on a series name all other series which are currently displayed will fade out to a gray color while the selected series is left unchanged and thereby making it selected stand out. In addition to this the selected series's tab will be selected in the information box, this will be further explained in section 7.3. If the user then clicks on a new series that series will be selected. However if it clicks on a series which is currently selected the other series will regain their color and the overview tab will be displayed in the information box.

The option to add a zoom function to the graph was also discussed which would allow the user to choose what time span should be displayed in the chart. The functionality for this was implemented but the option was removed from the final version since the frequency of the tweets were not as high as first expected. It was therefore deemed that the added buttons for controlling the functionality distracted more than the zooming option added to the total design.

The chart is designed so that the timeline will continuously scroll forward so that the current time always will be displayed. However with a lower amount of tweets on the subject coming in the chart has been set to show a larger time span than originally designed making this effect less visible than planed.

7.3 Information Box

In the initial concept development stage it was decided that the information box should be divided up into a tab system. The initial tab should contain a table showing the number of tweets and index (followed by an green/red arrow showing whether the index is going up or down) for each party as well as party leader. This will allow the user to directly get a clear picture of the current position of the parties.

In addition to the initial tab there will be a tab for each of the top 11 parties. Each so called party-tab will hold a pie chart showing the distribution of the party's tweets and a brief description of their current position. To the right of these it will also hold an area where party specific tweets will be displayed. They will be presented as a speech bubble which will be animated as if spoken from an image of a "Twitter bird." The speech bubble will be colored green, red or orange depending on whether it is respectively positive, negative or neutral in its sentiment. Three color coded buttons were also added in order to allow the user to choose if they only want to see positive, negative or neutral tweets for that party.

As previously mentioned in the line chart section the corresponding tabs will be selected when the user clicks on a series name in the legend of the line chart. The reverse also works, i.e. that when a party's tab is selected in the information box that party will be highlighted in the line chart. In the case that the selected tab's party isn't selected to be shown in the line chart all current series loose their colors and the selected party's is added.

In the case that a user clicks on a party's leader in the legend of the line chart that party’s tab is selected in the information box.

7.4 Statistics page

During the process of developing the concept further we realized that we had access to a lot of statistical information which we felt would be interesting to display but didn’t fit in on the page. It was therefore decided to add an extra page holding this information. This page was just to present the information and not as interactive and dynamic as the main page. The first section of the page was to hold a summary of the parties as well as the two main political blocks. This was then to be followed by a section for each of the parties.

The summary section consists of several different elements the first of which is the same table as the one in the information box on the main page. This is followed by a series of bar charts displaying the distribution of when Twitter users are the most active.

Chapter 8

Implementation

When implementing a website there are two important aspect to take into consideration. First the website needs to be accessible from all the major web-browsers, which demands that the WC3 standard is followed. Sadly, even though all WC3 demands are fulfilled it is still no guarantee that the site will work perfectly. The reasons is that web-browser sometimes brake the rules as well, especially Internet explorer. Secondly the site must be optimized so that even older computers and less powerful systems can view the page without to much delays. To assure these requirements to be fulfilled the implementation was performed in an iterative pattern. After each implementation session the site was validated through WC3 to minimize errors and warnings.

8.1 The Prototype

In the very early state of the project a prototype page was created to get a better understanding on how the the different components would look and function. The prototype consisted of adding one of the line-chart examples from FLOT's website and adding a text field to act as the tweet flow. The tweet flow and graph was both updated in regular time periods with nonsense text and values to make the site more living. The prototype helped creating a more realistic site that were useful during user-evaluations.

8.2 Main Application

As mentioned many times before, the application consisted of three main components:

- The tweet flow
- The line chart
- The information box

All of these components were design to be able to stand alone. They however all rely on an initialization function as well as some other pieces of code in order to parse the tweets for example. Due to the time constraints on the project there are here some pieces of code which can seem a bit redundant. An example of this can be seen in the number of functions used to convert text strings between different encoding standards. This was partly caused by the fact that several different languages have been used which implement different encodings. These can be reduced and homogenized given time but as previously stated it was deemed more important to have a fully functional product than perfect code.

8.3 Optimization

For optimization purposes all graphics were scaled to right size and format to unburden the site scaling and unnecessary loading times. The intention is also to implement a REST-API which function is to make sure polling the server for data is as minimal as possible. During the research phase of the project the intention was to implement a push server that sends data to the site directly as the server receives it to make sure data transfers is kept to a minimal. Saplo however recommended the polling structure, they insured that if designed correctly it would perform as well or better than the push structure. Anther optimization for the site was that all JavaScript was put in the very end of each HTML-page to make sure the site loaded directly without having to load the script.

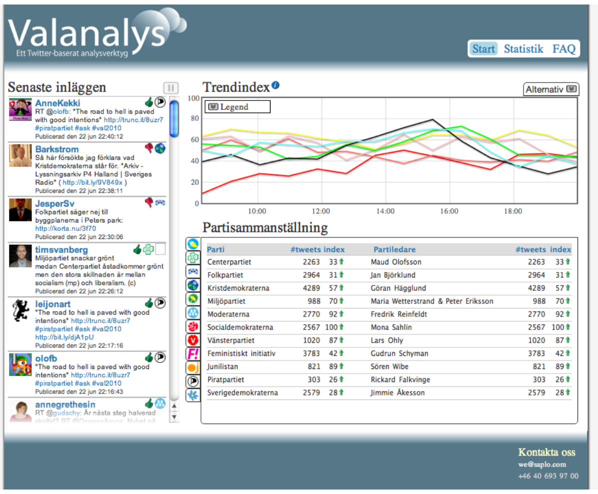
Chapter 9

Results

Saplo had an initial vision for the application with two of our three main components, the line chart and tweet flow. Saplo was however open for suggestions since achieving as good result as possible was in everyones best interest. Below is described the result from the different components and some other added functionality on the site.

9.1 Interface and Visual Results

The following section will describe the interface and visual result of the website. Before going into detail, the image below shows how the sites first page looks and how the different components are put together.



Senaste inläggen

AnneKekki
RT @olofb: "The road to hell is paved with good intentions" http://trunc.it/8uzr7 #piratpartiet #ask #val2010
Publicerad den 22 jun 22:40:12

Barkstrom
Så här försökte jag förklara vad Kristdemokraterna står för. "Arkiv - Lyssningsarkiv P4 Halland | Sveriges Radio" (http://bit.ly/9V849x)
Publicerad den 22 jun 22:38:11

JesperSv
Folkpartiet säger nej till byggplanerna i Peters park: http://korta.nu/3f70
Publicerad den 22 jun 22:30:06

timsvanberg
Miljöpartiet snackar grönt medan Centerpartiet åstadkommer grönt men den stora skillnaden är mellan socialism (mp) och liberalism. (c)
Publicerad den 22 jun 22:26:12

leijonart
"The road to hell is paved with good intentions" http://trunc.it/8uzr7 #piratpartiet #ask #val2010 http://bit.ly/djA1pU
Publicerad den 22 jun 22:17:16

olofb
"The road to hell is paved with good intentions" http://trunc.it/8uzr7 #piratpartiet #ask #val2010
Publicerad den 22 jun 22:16:43

annegrethesin
RT @gudschy: Är nästa steg halverad

Partisammanställning

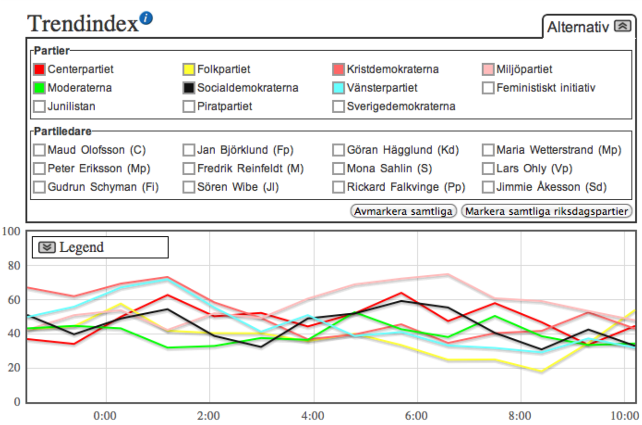
Parti	#tweets	index	Partiledare	#tweets	index
Centerpartiet	2263	33 ↑	Maud Olofsson	2263	33 ↑
Folkpartiet	2964	31 ↑	Jan Björklund	2964	31 ↑
Kristdemokraterna	4289	57 ↑	Göran Hägglund	4289	57 ↑
Miljöpartiet	988	70 ↑	Maria Wetterstrand & Peter Eriksson	988	70 ↑
Moderaterna	2770	92 ↑	Fredrik Reinfeldt	2770	92 ↑
Socialdemokraterna	2567	100 ↑	Mona Sahlin	2567	100 ↑
Vänsterpartiet	1020	87 ↑	Lars Ohly	1020	87 ↑
Feministiskt initiativ	3783	42 ↑	Gudrun Schyman	3783	42 ↑
Junilistan	821	89 ↑	Sören Wibe	821	89 ↑
Piratpartiet	303	26 ↑	Rickard Falkvinge	303	26 ↑
Sverigedemokraterna	2579	28 ↑	Jimmie Åkesson	2579	28 ↑

Figure 9.1. The start page

9.1.1 The Line Chart

As mentioned, Saplo wanted some kind of line chart that could display the trend indexes which are calculated from the incoming tweets. The tweets which contains keywords, like the name of the political party or the name of a political leader, and a trend index to plot the trend on the chart in real-time.

The framework Flot was used as a foundation for this line chart, however changes needed to be made to fit our requirements. The Flot graph was modified so that the graph is updated in frequent time intervals, making the graph come "alive". Our initial estimation was that the graph would constantly be moving. This however doesn't seems to be the case due to the yet low incoming tweet rate on the site, the graph is hardly moving at all but the functionality is implemented for a much faster paced graph. The second change that needed to be done was changing the time values on the x-axis in the graph to accurate dates for our time region. This was achieved by modifying the incoming tweet's timestamps to the correct format for the Flot graph. Thirdly, the legend on the Flot needed to be modified to better suit our site. When hovering over the graph with the mouse the legend will both display the elements names and their trend index for a specific time.



Partier

🟥 Centerpartiet 🟨 Folkpartiet 🟥 Kristdemokraterna 🟥 Miljöpartiet
🟩 Moderaterna ⬛ Socialdemokraterna 🟦 Vänsterpartiet □ Feministiskt initiativ
□ Junilistan □ Piratpartiet □ Sverigedemokraterna

Partiledare

□ Maud Olofsson (C) □ Jan Björklund (Fp) □ Göran Hägglund (Kd) □ Maria Wetterstrand (Mp)
□ Peter Eriksson (Mp) □ Fredrik Reinfeldt (M) □ Mona Sahlin (S) □ Lars Ohly (Vp)
□ Gudrun Schyman (Fi) □ Sören Wibe (Jl) □ Rickard Falkvinge (Pp) □ Jimmie Åkesson (Sd)

Figure 9.2. Line chart expanded

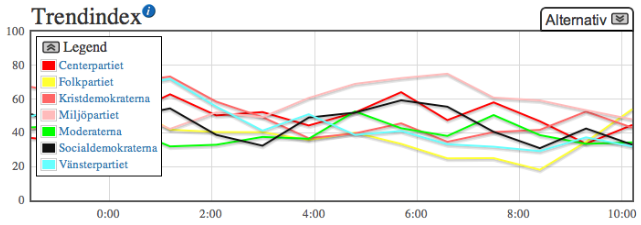
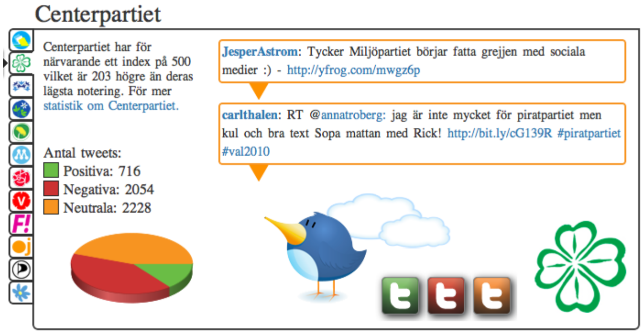


Figure 9.3. Line chart closed

9.1.2 Information Box

The information box was not fully developed in the initial concept design for the project, its design grew during the process of implementation. Its purpose is to work as a compliment to the line chart, where the trend indexes can be displayed and where each party has its own summary. In the party summary section of the information box the user can view old tweets and sort them by their sentiment. This is presented by a twitter bird "saying" the tweet in form of a speech bubble to make it a bit more playful. There is also a pie chart showing the distribution of all the tweets concerning the political party, how many that are positive, negative and that has no opinion at all.



Centerpartiet

Centerpartiet har för närvarande ett index på 500 vilket är 203 högre än deras lägsta notering. För mer statistik om Centerpartiet.

Antal tweets:

- 🟩 Positiva: 716
- 🟥 Negativa: 2054
- 🟧 Neutrala: 2228

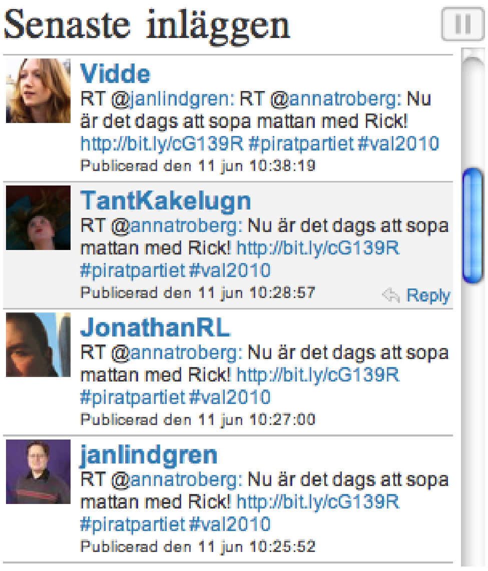
JesperAstrom: Tycker Miljöpartiet börjar fatta grejjen med sociala medier :) - http://yfrog.com/mwgz6p

carlthalen: RT @annatroberg: jag är inte mycket för piratpartiet men kul och bra text Sopa mattan med Rick! http://bit.ly/cG139R #piratpartiet #val2010

Figure 9.4. Informationbox

9.1.3 Tweet Flow

The tweet flow responsibility is to present all messages from Twitter in real-time. The flow is placed to the far left on the site, making it the logical order for the user to first read the tweets and when watching the other elements on the site. Each message contains information about the person that posted the message: the avatar image, name of the user and the exact time when the tweet was posted on Twitter. To visualize which political party the message was concerning we added a method to automatically add the correct party symbol beside the users name. A green thumb alternatively a red thumb is presented next to the party symbol representing if the message is positive or not. If the tweet is neutral, no thumb symbol will be shown. When the user hovers over the message a reply button becomes more visible than before, down in the bottom right corner.



Senaste inläggen

Vidde
RT @janlindgren: RT @annatroberg: Nu är det dags att sopa mattan med Rick! http://bit.ly/cG139R #piratpartiet #val2010
Publicerad den 11 jun 10:38:19

TantKakelugn
RT @annatroberg: Nu är det dags att sopa mattan med Rick! http://bit.ly/cG139R #piratpartiet #val2010
Publicerad den 11 jun 10:28:57
Reply

JonathanRL
RT @annatroberg: Nu är det dags att sopa mattan med Rick! http://bit.ly/cG139R #piratpartiet #val2010
Publicerad den 11 jun 10:27:00

janlindgren
RT @annatroberg: Nu är det dags att sopa mattan med Rick! http://bit.ly/cG139R #piratpartiet #val2010
Publicerad den 11 jun 10:25:52

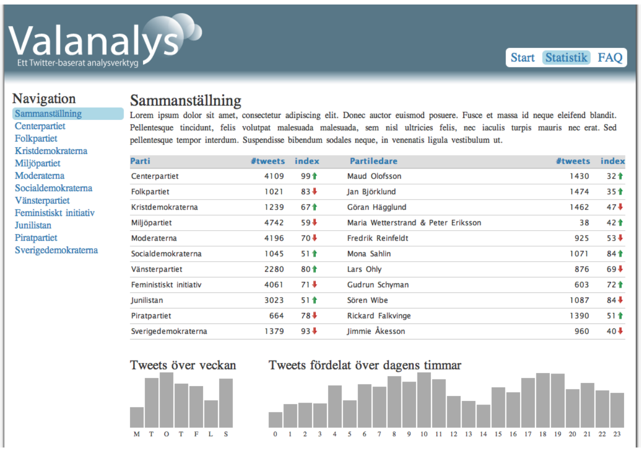
Figure 9.5. The tweet-flow

9.1.4 The Statistics Page

The statistics page was something that we added rather late in the project, because we wanted the possibility to add more detailed information about the different political parties on the site.

In the top of the statistics page we have placed an overview and summary of all the political parties together with their trend indexes in a table identical to the table on the main site. There are general statistics on how many tweets the site has gathered during different time periods; the users can view how many tweets that has been registered for each day. This gives statistics about which week days as well as hour of the day that people are most active at writing about politics. There is also a section in the overview for the two major political alliances: the "Alliansen" and the "Rödgröna". Here the user can view the collected trend analysis for the different alliances and how many tweets each alliances has.

Each of the different parties has its own section on the statistics page, and all of them have the exact same layout for the information. Just like on the main page, the user can see how many tweets a party has as well as their sentiment distribution. But there is much more, the user can also view the frequency of the tweets during the hours of the day, the current trend index together with the trend graph. The user can also view a collection of old tweets relevant to the political party.



Sammanställning

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec auctor euismod posuere. Fusce et massa id neque eleifend blandit. Pellentesque tincidunt, felis volutpat malesuada malesuada, sem nisl ultricies felis, nec iaculis turpis mauris nec erat. Sed pellentesque tempor interdum. Suspendisse bibendum sodales neque, in venenatis ligula vestibulum ut.

Parti	#tweets	index	Partiledare	#tweets	index
Centerpartiet	4109	99↑	Maud Olofsson	1430	32↑
Folkpartiet	1021	83↓	Jan Björklund	1474	35↑
Kristdemokraterna	1239	67↑	Göran Hägglund	1462	47↓
Miljöpartiet	4742	59↓	Maria Wetterstrand & Peter Eriksson	38	42↑
Moderaterna	4196	70↓	Fredrik Reinfeldt	925	53↓
Socialdemokraterna	1045	51↑	Mona Sahlin	1071	84↑
Vänsterpartiet	2280	80↑	Lars Ohly	876	69↓
Feministiskt initiativ	4061	71↓	Gudrun Schyman	603	72↑
Junilistan	3023	51↑	Sören Wibe	1087	84↓
Piratpartiet	664	78↓	Rickard Falkvinge	1390	51↑
Sverigedemokraterna	1379	93↓	Jimmie Åkesson	960	40↓

Figure 9.6. The statistic page

9.1.5 FAQ

This page was added to present some information about how the site works, why it has been created and give Saplo space to present themselves. The page has a very basic layout with questions that we gathered from our user evaluations, which we try to answer together with Saplo.

9.2 Back-End Results

In the section above we have described the visual result for our work, the result that the users will come in contact with. However, there are also a lot of technical results we have achieved during this project to make the site work as it should. Below is a short summary of what we have done, from collecting the information from Twitter to visually presenting it on our site.

9.2.1 Handling Data Transfers

For this site to work we needed to gather information from social networks, we chose to limit ourselves to Twitter for this project due to our time constraint. Saplo stood for most of the back-end structure development, which consists of gathering the information from Twitter by searching for specific keywords concerning the Swedish election. They sent the incoming tweets to their analysis program to create a trend index. When the analysis is completed the information is sent back to the web server so that it in turn can be requested by the client. In order to have a structured order of the information we chose to order the data by using JSON format.

9.2.2 Component Based Development

As mentioned in section 4, we aimed to create a site that contains three different elements: Tweet flow, line chart and information box. All these components were implemented and fully functional although they did not become as independent as planed. These components need some code that combines them in order to function correctly. To make these components work alone, that common code needs to be modified to only serve the components separately.

Chapter 10

Discussion and Conclusions

Developing Valanalys.se has truly been an exciting process which required time, knowledge and effort. The site required us to perform a lot of software research to find the most suitable frameworks to fulfill Saplo's vision. Although the research was necessary to get the best result, it would be much more exciting to spend time researching interfaces and implementing them. Overall we are very pleased with the result of our work, a fully functional website that searches Twitter for relevant posts concerning the election, analyzes them and presents them in real-time. There are however some areas we would like to improve, they are discussed in the section below.

10.1 Areas of Improvement

We were able to implement all the core functionally and the final result is something that we are proud of. We even added the statistics section that gives people whom are interested in an even more detailed view than that which is on our main page. Even though we are proud of the result there are some areas that could be improved.

10.1.1 Social Media and News Article Usage

As mentioned in the section 1.4 of the report, this master thesis only focused on the social media Twitter. It would be very interesting to expand to other media as well as newspapers. The newspapers could be a compliment to the tweet stream by also analyzing that type of media. Due to the longer texts in articles Saplo could give more accurate trends for the different political parties and leaders. The news articles could also be used in another matter, they could be linked together with the timeline in the line chart. This way a user could select to read relevant news articles connected to specific times in the chart. The result would lead to a more complete experience.

10.1.2 Line Chart Improvement

The line chart is an area that can be improved in many ways. One thing which we think would be very interesting for the user is to allow them to create his/her own charts by merging different political parties. The outcome would be a new chart that represents the trend for all the parties together, this would be very useful when comparing for instance the two different political alliances, "Alliansen" and the "Rödgröna". By comparing them the user could for instance merge another party to see if that party would play an important part for the collective positive trend for the parties.

Another very interesting addition to the line chart is to use multidimensional scaling or MDS for short [21]. It is commonly used in information visualization for comparing similarities and differences in data. This can result in a three dimensional graph where the political parties and the leaders can be compared and grouped according to their similarities and differences. To get a good result with MDS it would be necessary to gather more information than we do from Twitter. This type of grouping would require the use of news articles and similar types of larger sources to help Saplo create an accurate opinion on the different parties and political leaders.

Both in the "normal" line chart and in the MDS, the possibility to go deeper and analyse how the different parties and leaders affect each other would also be an intriguing feature. This could interactively show correlations between parties and party leaders. But then again it would require much more information than Twitter can provide.

10.1.3 Component Based Development

The intention was to develop every component on the site to be stand-alone. This feature would be invisible for the users but it would be very convenient for Saplo, both for further development, maintainability and if they want to separate components for other products or systems. However this design became compromised in the near end of the project due to lack of time. For the time being it was more important to get the product complete than having a perfect code structure. The question is whether this inconvenience could have been prevented by better planing? Maybe, but it is very difficult to tell. When Saplo initially set its priorities for this project it was to deliver a complete and functional product within a very limited time frame. With this knowledge we tried to plan the project and eventual problems to the best of our ability. There is however always a balancing act in how much planing can be done before one can start to implement. So even if a project is planed into every detail there are some problems which can’t be foreseen during the planing phase.

Even though the code structure isn’t as perfect as we would have hoped in the beginning the result is yet pleasing but the code structure could absolutely be improved for the reasons previously mentioned.

10.1.4 Interface, Usability and Alternative Views

Coming from an interaction design background we were eager to apply our knowledge in that area in this project. The project could easily have been a pure interaction design master thesis, designing smart and easy to use interfaces due to its unique and new thinking nature. However the project required a much broader software development process which of course required a lot of the projects time. So unfortunately we couldn't spend as much time on the interface design and user evaluations as we wanted. This project really had the potential of being a pure interaction design master thesis, where all the time could be spent on improving and analyzing different interface solutions. The alternative views that are mentioned in the section with the same name are an excellent example of exciting ideas that would be very fun to explore further.

Of all our different ideas for alternative views, SapGarden was by our opinion the most innovative and interesting of them all. We believe that this interface could be very visually appealing and able to contain a lot of information in a fun way that would evoke peoples interest.

10.1.5 Multi-Platform Support

The decision to use JavaScript before anything else was mainly based on that JavaScript works very well on many different platforms, for example the mobile platforms. When developing this site we had other platforms in mind but couldn't focus on optimizing for them. Most of the time we had reserved for optimization went to make sure the site worked in different web browsers. Internet Explorer can be quite tedious due to its peculiar nature of not keeping within the Internet standards. For instance the site today works in all major web browsers but could visually differ a tiny bit.

The next step in this section for future development would be to optimize the site for the iPhone due to it is popularity. It would also be an interesting choice to develop an entire application for the iPhone, optimizing for its resolution and the interface limitations which comes along with it.

10.2 Final Words

It has been a very a fun and educational journey developing Valanalys.se, both from an interaction design point of view and a more traditional software development aspect. We hope that the site will attract a lot of people and that they will have good use of the sites functionality. The website has a great potential for further development and added functionality. Our hope is that we have met or exceeded Saplo's vision of the site, that they will enjoy it and that it helps them with their marketing goal.

Bibliography

[1] Abode. About flash. Retrieved: May 15 2010, http://www.adobe.com/products/flashplayer/.

[2] Helen Ahlbom and Charlotta von Schultz. It-bolag avslöjar politiska spelet, apr 2010. Retrieved: May 26 2010, http://www.nyteknik.se/nyheter/it_telekom/allmant/article760253.ece.

[3] Aitellu. About aitellu. Retrieved: May 27 2010, http://www.aitellu.com/Aitellu/about.html.

[4] Martin Bryant. Twitter geo-fail? only 0.23% of tweets geotagged, jan 2010. Retrieved: May 24 2010, http://thenextweb.com/2010/01/15/twitter-geofail-023-tweets-geotagged/.

[5] Microsoft Corporation. About silverlight. Retrieved: May 14 2010, http://www.silverlight.net/.

[6] Oliver Dan. Why html5 isn't a flash killer. Retrieved: Mar 20 2010, http://www.techradar.com/news/software/applications/why-html5-isn-t-a-flash-killer-674741.

[7] Emprisecorporation. ejschart. Retrieved: Mar 07 2010, http://www.ejschart.com/.

[8] Matthew Fraser and Somumitra Dutta. Obama's win means future elections must be fought online, nov 2008. Retrieved: Jun 11 2010, http://www.guardian.co.uk/technology/2008/nov/07/barackobama-uselections2008.

[9] Jesse James Garrett. Ajax: A new approach to web applications, feb 2005. Retrieved: May 12 2010, http://www.adaptivepath.com/ideas/essays/archives/000385.php.

[10] Google. google chart. Retrieved: Jun 11 2010, http://code.google.com/apis/charttools.

[11] http://www.makthavare.se/. About twixdagen. Retrieved: May 28 2010, http://twixdagen.se/om/.

[12] Ideo.com. Ideo. Retrieved: Jun 02 2010, http://www.ideo.com.

[13] json.org. json. Retrieved: May 27 2010, http://www.json.org/.

[14] Mozilla Developer Network. About javascript. Retrieved: Mar 27 2010, https://developer.mozilla.org/en/About_JavaScript.

[15] olau@iola.dk. Flot. Retrieved: Mar 07 2010, http://code.google.com/p/flot/.

[16] php.net. Php. Retrieved: Mar 05 2010, http://www.php.net/.

[17] plotkit.com. Plotkit. Retrieved: Mar 05 2010, http://www.liquidx.net/plotkit.

[18] Erick Schofeld. Twitter turns on the firehose for real-time search startups@ONLINE. Retrieved: Apr 06 2010, http://techcrunch.com/2010/03/01/twitter-firehose-realtime-search-startups.

[19] Nigel Slack, Stuart Chambers, and Robert Johnston. *Operation Management*, pages 507–508. Prentice Hall, 2007.

[20] Nigel Slack, Stuart Chambers, and Robert Johnston. *Operation Management*, pages 306–507. Prentice Hall, 2007.

[21] Statsoft. Multidimensional scaling. Retrieved: May 31 2010, http://www.statsoft.com/textbook/multidimensional-scaling/.

[22] trends.builtwith.com. Websites using jquery. Retrieved: May 05 2010, http://trends.builtwith.com/javascript/JQuery.

[23] Twingly. About twingly. Retrieved: May 27 2010, http://www.twingly.com/about.

[24] Twingly. Twingly blogstream. Retrieved: May 27 2010, http://www.twingly.com/enterprise.

[25] Twitter. About twitter. Retrieved: Feb 07 2010, http://twitter.com/about.

[26] W3Schools. Asp .net. Retrieved: May 05 2010, http://www.w3schools.com/aspnet/aspnet_vsasp.asp.

[27] Wikipedia. ajax cross domain. Retrieved: May 05 2010, http://en.wikipedia.org/wiki/Same_origin_policy.

[28] Wikipedia. Same origin policy. Retrieved: May 05 2010, http://en.wikipedia.org/wiki/Same_origin_policy.

[29] www.conversition.com. About tweetfeel. Retrieved: May 28 2010, http://www.tweetfeel.com/faq.php.